DOCUMENT RESUME

ED 457 175 TM 033 265

TITLE	The Development of STAR Early Literacy. Report.
INSTITUTION	School Renaissance Inst., Inc., Madison, WI.
PUB DATE	2001-03-00
NOTE	14p.; School Renaissance Institute, Inc. became Renaissance Learning, Inc.
AVAILABLE FROM	Renaissance Learning, Educational Research Department, P.O. Box 45016, Madison, WI 53744-5016. Tel: 800-200-4848 (Toll Free); Web site: http://www.SchoolRenaissance.com.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Adaptive Testing; *Beginning Reading; *Computer Assisted Testing; Diagnostic Tests; Early Childhood Education; *Emergent Literacy; Reading Readiness; Screening Tests; *Test Construction; *Validity

ABSTRACT

This report describes the development and testing of a computerized early literacy diagnostic assessment for students in prekindergarten to grade 3 that can measure skills across a variety of preliteracy and reading domains. The STAR Early Literacy assessment was developed by a team of more than 50 people, including literacy experts, psychometricians, item developers, graphic artists, audio experts, and software engineers. More than 50,000 students in 400 schools across the United States participated in STAR Early Literacy assessment development. STAR Early Literacy is a computer-adaptive assessment and database that helps educators identify a student's command of phonemic awareness, phonics, general readiness, graphophonemic knowledge, comprehension, structural analysis, and vocabulary in approximately 10 minutes. It is designed to be a low-stakes assessment that gives teachers a tool to align instruction to the needs of each student even though students require little teacher assistance while taking the assessment. Details are provided about content specification, item development, software and user interface design features. Also reported are the prototype research study involving 1,500 children from grades prekindergarten through 2 and the item calibration study involving 32,257 students in 308 schools. Other research data are being collected with a pilot adaptive version of STAR Early Literacy. Data to date indicate that the STAR Early Literacy diagnostic assessment meets the need for an accurate, inexpensive tool to measure pre-reading skills and early literacy skills in seven domains. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 457 175

REPORT from the School Renaissance Institute

March 2001

The Development of STAR Early Literacy

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Vander Galien

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

TM033265

BEST COPY AVAILABLE

REPORT from the School Renaissance Institute

March 2001

The Development of STAR Early Literacy

Introduction and Design Goals

The development of a computerized, early literacy diagnostic assessment for students in pre-K to grade 3 that can measure skills across a variety of pre-literacy and reading domains has been much awaited. According to the National Research Council's study *Preventing Reading Difficulties in Young Children:*

> Much has been learned about which particular differences among preschoolers and kindergartners are most prognostic of early reading outcomes, and these findings, in turn, have enabled more effective programs of intervention. However, the array of instruments currently used to measure such differences are time consuming and costly to administer, even as they are mutually redundant and collectively incomplete with respect to the range of knowledge and sensitivities on which reading growth, including longer-term reading growth, depends.[1]

To meet the need for such an assessment tool, Renaissance Learning started development of STAR Early Literacy diagnostic reading assessment three years ago. STAR Early Literacy was developed by a team of over 50 people including literacy experts, psychometricians, item developers, graphic artists, audio experts, and software engineers. Over 50,000 students in 450 schools nationwide participated in STAR Early Literacy development.

The design goals for STAR Early Literacy were as follows:

1. To develop a valid and reliable criterion-referenced assessment of student abilities in the pre-reading skills most important to later reading success.
2. To administer this assessment automatically via computer.
3. To enable assessments to be completed in 10 minutes or less.
4. To provide the ability to administer the assessment multiple times during a year for progress tracking.
5. To provide individual and class reporting.
6. To significantly reduce the cost compared to traditional paper assessments.

Description of STAR Early Literacy

STAR Early Literacy is a computer-adaptive assessment and database that helps educators identify a student's command of phonemic awareness, phonics, general readiness, graphophonemic knowledge, comprehension, structural analysis, and vocabulary in approximately 10 minutes. STAR Early Literacy was designed to be used as a low-stakes assessment to provide teachers a tool to align instruction to the needs of each student and accelerate literacy development. The design is consistent with well-recognized principles of literacy development, including the *Principles and Recommendations for Early Childhood Assessments*[2] produced by the National Educational Goals Panel, and the federal Reading Excellence Act.

STAR Early Literacy employs multimedia and computer-adaptive technology to ensure that students require minimal teacher assistance while taking the assessment. Questions continually adjust in difficulty based on a student's previous response, thereby reducing frustration. When help is needed, audio alerts prompt students to ask for assistance. The software's graphics, clear audio instructions, and other features enable students to take the assessment independently, while assuring a comfortable and enjoyable experience.

STAR Early Literacy provides educators with immediate, accurate, and reliable feedback on students' literacy progress. As a result, educators are able to intervene sooner and provide students with effective instruction during the most critical years of their literacy development. STAR Early Literacy's detailed reports help educators identify student literacy development levels, assess and demonstrate progress, determine instructional focus, and strengthen parent communications. The four STAR Early Literacy sample reports which follow show the information that is provided:

---

[1] C.E. Snow, M.S. Burns, and P. Griffin, eds. *Preventing Reading Difficulties in Young Children.* (Washington, DC: National Academy Press, 1998: p. 336).

[2] L. Shepard, S.L. Kagan, and E. Wurtz, eds. *Principles and Recommendations for Early Childhood Assessments.* (Washington, DC: National Educational Goals Panel, 1998).

Growth Report

STAR Early Literacy™ : Wednesday, November 4, 2001
Reporting Period: 8/11/01-11/04/01 (Fall 01)

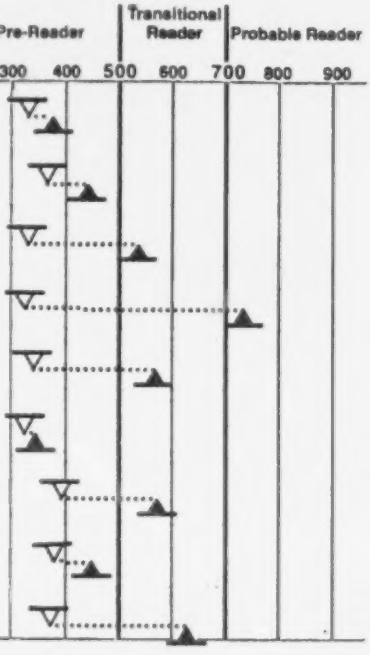
Mayfield Elementary

Sorted By: Student Name

Class: Kinder A

Teacher: Kevin Wright

Student Name	Age (yrs)		Literacy Domain Scores GR	GK	PA	PH	SA	VO	CO	Scaled Score
Clark, William J.	4.7	09/01/01	65	55	50	45	40	50	45	325
	5.0	11/01/01	78	70	65	53	44	53	45	375
Garcia, Maria D.	4.8	09/01/01	60	40	35	30	25	20	25	363
	5.1	11/01/01	74	52	43	35	27	21	25	425
Jackson, Betty M.	4.9	09/01/01	75	60	50	45	40	35	40	327
	5.2	11/01/01	98	76	63	54	43	37	40	539
Moore, Christopher	4.7	09/01/01	60	40	35	30	25	20	25	322
	5.0	11/01/01	73	48	43	35	26	21	25	728
Perez, Charles U.	4.8	09/01/01	70	55	45	40	35	30	35	332
	5.1	11/01/01	87	70	55	47	38	31	36	562
Thompson, Kevin	4.8	09/01/01	65	50	40	35	30	25	30	327
	5.1	11/01/01	80	60	48	39	32	25	31	334
Walker, Mark V.	5.0	09/01/01	65	50	40	35	30	25	30	391
	5.5	11/01/01	79	61	51	40	32	26	30	576
Webster, Andrew	5.1	09/01/01	65	50	40	35	30	25	30	378
	5.2	11/01/01	84	63	51	40	32	25	31	451
Willis, Ricardo M.	5.2	09/01/01	75	60	50	45	40	35	40	372
	4.5	11/01/01	95	73	61	51	44	37	41	618



Diagnostic Report

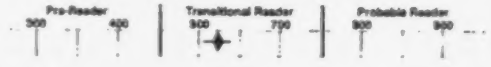
STAR Early Literacy™ : Wednesday, November 4, 2001
Reporting Period: 8/11/01-11/04/01 (Fall 01)

Mayfield Elementary

Clark, William J.
Grade: K

Teacher: Kevin Wright
Class: Kinder A

Age (yrs)	Last Test	Scaled Score
4.7	09/02/00	539



Strengths and Weaknesses by Skill Score

	< 25	25 - 49	50 - 75	> 75
General Readiness	Completing sequences	Comparing word length (written) Differentiating letters Differentiating words from letters Differentiating shapes	Recognizing position words Matching numbers and objects Identifying word boundaries	Differentiating word pairs
Grapheme/Phoneme Knowledge		Recognizing letter sounds Using alphabetical order	Naming letters	Matching upper and lower case letters Recognizing alphabetic sequence
Phonemic Awareness	Blending word parts Blending phonemes	Discriminating sounds Identifying missing sounds		Identifying rhyming words Comparing word length (oral)
Phonics	Replacing beginning and ending consonants Replacing vowels Identifying consonant blends Identifying consonant digraphs	Matching sounds within word families	Matching and recognizing short vowel sounds Identifying ending consonant sounds Identifying medial short vowels	Matching and recognizing long vowel sounds Identifying beginning consonant sounds Identifying medial long vowels Substituting consonant sounds
Structural Analysis		Finding words Building words Identifying compound words		
Vocabulary			Recognizing synonyms Recognizing antonyms	Matching words and pictures
Comprehension	Reading and understanding words Reading and completing sentences Reading and understanding paragraphs			

BEST COPY AVAILABLE

Score Distribution Report

STAR Early Literacy™ : Wednesday, November 4, 2001
Reporting Period: 8/11/01-11/04/01 (Fall 01)

Mayfield Elementary

Class: Kinder A **Teacher: Kevin Wright**

General Readiness

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Comparing word length (written)	4	5	4	3
Recognizing position words	4	5	4	3
Differentiating letters	4	5	4	3
Differentiating words from letters	4	5	4	3
Matching numbers and objects	4	5	4	3
Differentiating word pairs	4	5	4	2
Identifying word boundaries	4	5	4	2
Differentiating shapes	4	5	4	3
Completing sequences	4	5	4	3
Completing sequences	4	5	4	3
Average	4	5	4	3

Graphophonemic Knowledge

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Matching upper and lower case letters	4	5	4	3
Recognizing alphabetic sequence	4	5	4	3
Naming letters	4	5	4	3
Recognizing letter sounds	4	5	4	3
Using alphabetical order	4	5	4	2
Average	4	5	4	3

Phonics

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Matching and recognizing long vowel sounds	4	5	4	3
Matching and recognizing short vowel sounds	4	5	4	3
Identifying beginning consonant sounds	4	5	4	3
Identifying ending consonant sounds	4	5	4	3
Replacing beginning and ending consonants	4	5	4	2
Replacing vowels	4	5	4	3
Identifying medial short vowels	4	5	4	3
Identifying medial long vowels	4	5	4	3
Matching sounds within word families	4	5	4	3
Identifying consonant blends	4	5	4	3
Identifying consonant digraphs	4	5	4	3
Substituting consonant sounds	4	5	4	3
Average	4	5	4	3

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Identifying rhyming words	4	5	4	3
Blending word parts	4	5	4	3
Blending phonemes	4	5	4	3
Discriminating sounds	4	5	4	3
Comparing word length (oral)	4	5	4	3
Identifying missing sounds	4	5	4	2
Average	4	5	4	3

Structural Analysis

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Word finding	4	5	4	3
Word building	4	5	4	3
Identifying compound words	4	5	4	3
Average	4	5	4	3

Vocabulary

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Matching words and pictures	4	5	4	3
Recognizing synonyms	4	5	4	3
Recognizing antonyms	4	5	4	3
Average	4	5	4	3

Comprehension

	Number of Students in Class with Skill Scores:			
	< 25	25-49	50-75	> 75
Reading and understanding words	4	5	4	3
Reading and completing sentences	4	5	4	3
Reading and understanding paragraphs	4	5	4	3
Average	4	5	4	3

Parent Report

STAR Early Literacy™ : Wednesday, November 4, 2001
Test Date: 11/01/01

Mayfield Elementary

Wren, Thomas Q. **Teacher: Kevin Wright**
Grade: K **Class: Kinder A**

Dear Parent or Guardian:

Your child has just taken a STAR Early Literacy™ assessment on the computer. STAR Early Literacy measures your child's proficiency in up to seven areas that are important in reading development. This report summarizes your child's scores on the assessment. As with any assessment, many factors can affect your child's scores. It is important to understand that these scores provide only one picture of how your child is doing in school.

Scaled Score: 375
The Scaled Score is the overall score that your child received on the STAR Early Literacy assessment. It is calculated based on both the difficulty of the questions and the number of correct responses. Scaled Scores in STAR Early Literacy range from 300 to 900 and span the grades Pre-K through 3.

Thomas obtained a Scaled Score of 375. This is an increase of 50 from the Scaled Score of 325 that Thomas obtained on the first taking of the assessment. Scaled Scores relate to three developmental stages: Pre-Reader (300 - 499), Transitional Reader (500 - 699), and Probable Reader (700 - 900). A Scaled Score of 375 means that Thomas is at the Pre-Reader stage.

Date Tested	Scaled Score	Pre-Reader 300 400	Transitional Reader 500 700	Probable Reader 800 900
09/01/01	325	▽		
11/01/01	375	▲		

▽ Initial Test Scaled Score
▲ Last Test Scaled Score

BEST COPY AVAILABLE

STAR Early Literacy Development Process

Content Specification

Content development for STAR Early Literacy was driven by the design and intended usage of the test. The desired content had to meet certain criteria. First, it had to cover a range of difficulty broad enough to test students from pre-kindergarten through third grade. It also had to allow testing of remedial students in grades four and above. Second, the final collection of test items had to be large enough so students could test more than 10 times per year. Third, there had to be test items for assessing skills in 7 domains and 41 skill areas.

Extensive research into the pre-reading and reading skills necessary for later reading success revealed that STAR Early Literacy would need to cover the broad language arts areas of listening and reading. Proposed item content was grouped into the following seven domains, each considered essential to reading development:

1. General Readiness–Understanding of written word length, position words, words vs. letters, basic numeracy, word matching, word boundaries, shapes, and patterns.
2. Graphophonemic Knowledge–Understanding of letter names and sounds, alphabetic letter sequence, and alphabetical order.
3. Phonemic Awareness–Understanding of rhyming words, ability to blend word parts and phonemes (speech sounds), sound discrimination, oral word length, and ability to identify missing sounds.
4. Phonics–Understanding of long vowels, short vowels, beginning and ending consonants, consonant and vowel replacement, word families (onset and rime), consonant blends, clusters, and digraphs.
5. Comprehension–Ability to read and derive meaning from words, sentences, and paragraphs.
6. Structural Analysis–Ability to find words within other words, build words, and compound words.
7. Vocabulary–Identify high frequency words, synonyms, and antonyms.

An item blueprint was then constructed, detailing the individual skills, item types, and grade level distributions needed for each domain.

Item Development

During item development, every effort was made to avoid the use of stereotypes, potentially offensive language or characterizations, and descriptions of people or events that could be construed as being offensive, demeaning, patronizing, or otherwise insensitive. The editing process also included a strict sensitivity review of all items to address issues of gender and ethnic-group balance and fairness.

Once the test design was determined, individual test items were developed for tryout and calibration. A total of 2,991 items, comprised of 2,961 test items and 30 mouse training items, were developed according to the following specifications:

- **Simplicity**
  Items should directly address the domain and skill objective in the most straightforward manner possible. Evaluators should have no difficulty deducing the exact nature of the skill being assessed by the item. Instructions should be explicit and consistent from one item to the next.

- **Screen Layout**
  The testing screen should be comfortable for the student and teacher. Background colors should be unobtrusive and relatively muted. Text and graphics should stand out clearly against the background. The item background must be the same for all items on the test. Each item should consist of some combination of audio instructions, an on-screen prompt in the form of a cloze stem containing text or graphics, and two or three answer choices containing letters, words, graphics, and sound.

- **Text**
  For letter and word identification items, the type size should be relatively large, becoming smaller as grade level increases. The type size should be tied to items, so that it varies according to the developmental level of a student; in other words, easier items should have larger type than more difficult items because the difficulty will correspond roughly to grade placement.

  All STAR Early Literacy test items will be administered auditorily by the computer, so there should be no need for printed directions on-screen. For items that require on-screen directions, the type should be a serif font of appropriate size.

  Every effort should be made to use common words as the target and distracter words in test items.

  For phonemic awareness and phonics items, the 44 phonemes that make up the English language should be used. Phonemes should be depicted in recording scripts by one or more letters enclosed in a beginning and ending forward slash mark.

- **Graphics**
  Any art should be easily recognized by students. Color should be functional, as opposed to decorative, and lines should be as smooth as possible. For complex graphics, such as those needed for

listening comprehension, line drawings on a light background should be used. The size and placement of the graphics should be consistent throughout.

The art for correct answers and distracters should be consistent in order to avoid introducing an extraneous error source. Answer choices will primarily consist of graphics and text, but sound or animation occasionally will be needed. Art should be acceptable to a broad range of teachers, parents, and students, avoiding controversial or violent graphics of any kind.

- **Answer Options**
  As a general rule, items should have three answer choices. Only one of the choices should be the correct answer. Answer choices should be arranged horizontally.

  Distracters should be chosen to provide the most common errors in recognition, matching, and comprehension tasks.

  Words and artwork used in answer choices should be reused in no more than 10 percent of the items within a skill, a domain, or within the item bank as a whole.

- **Language and Pronunciation**
  Language should be used consistently throughout the assessment. Standard protocols should be established for item administration that reflect consistent instructions. For example, if an item stem is repeated twice, the same repetition should be used for all items of the same type. One exception to this rule is those situations where the same item type is used across grades, and one of the factors that changes is the level of instruction provided to the student.

  In phonemic awareness items, words should be segmented into phonemes, that is, divided into their individual sounds. As much as possible, the individual sounds should be preserved, and not distorted in any way. In the item instructions, individual phonemes will be enclosed by two forward slash marks.

  In the recording of item instructions and answer sounds, the audio segments should minimize the tendency to add a vowel sound after a consonant sound, especially for unvoiced consonants, such as /p/, /k/, and /t/. For example, /p/ should not be pronounced "puh". Instead, it should be spoken in a loud whisper and in a clipped manner.

  For voiced consonants that cannot be pronounced without a vowel sound, such as /b/ and /g/, the audio segments should keep the vowel sound as short as possible. For example, /g/, not "guh".

  Constituent consonants, such as /m/, /f/, and /n/, should not be followed by a vowel sound. They can, however, be extended slightly, as in "mmmmm", but not "muh".

  Short and long vowel sounds should be pronounced by simply lengthening the sound of the vowel. The long "a" sound, for example, should be pronounced "aaaaa".

Software and User Interface Design Features
The STAR Early Literacy user interface was designed to be simple and effective, allowing for a comfortable experience for the child. Prior to actual test administration, the child is given pretest instructions on-screen on how to use the mouse, how to use the <Listen> button (which repeats instructions), and how to select an answer. He is then led through a series of screens that check his ability to use the mouse and his understanding of instructions. The software closely tracks the child's responses and posts a graphical teacher alert if it detects that a child is struggling. STAR Early Literacy assessments are administered in the following three parts:

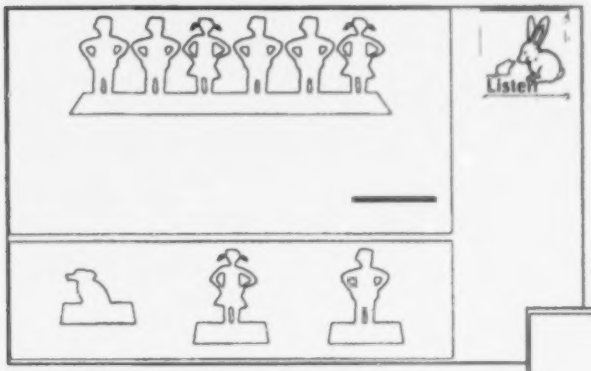
1. Mouse training–a series of mouse training items with a single answer choice and instructions that prompt the child to click on the object in the answer choice. The child needs to demonstrate a level of mouse proficiency in order for the practice test to begin.

2. Practice test items–a series of practice test items targeted at a level below that of the child's grade or age. Practice items have three answer choices and instructions that ask the child a pre-literacy question. The child needs to click the correct answer for three out of five practice items. If he does not, a teacher alert is posted on the screen. The teacher will be asked to assist the child in answering the practice items a second time.

3. Actual test items–a series of 25 test items targeted at the ability level of the child. These items have up to three answer choices and audio instructions, similar to the practice items. The test ends when th child has answered all of the test items.

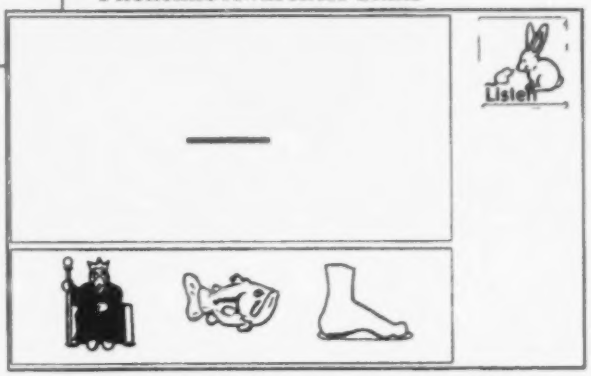
Seven sample STAR Early Literacy screenshots are shown on the following pages:

General Readiness Skills



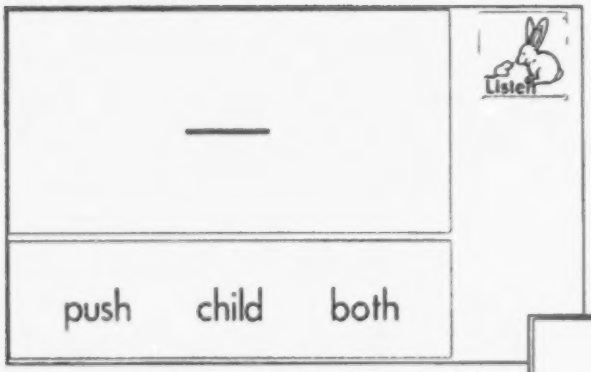
Look at the pattern.
Click on the picture that comes next.

Phonemic Awareness Skills



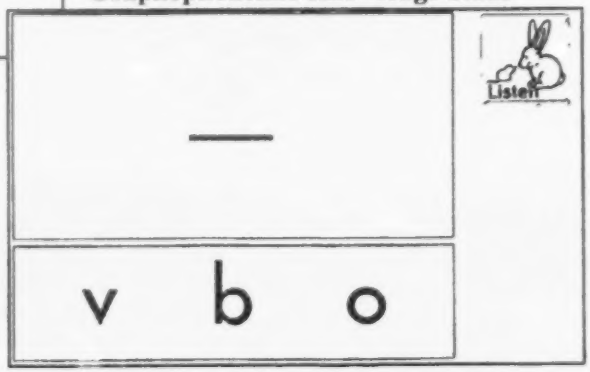
Listen carefully. The pictures are king, fish, foot.
Click on the picture that has a different beginning sound than the others... king, fish, foot.

Phonics Skills



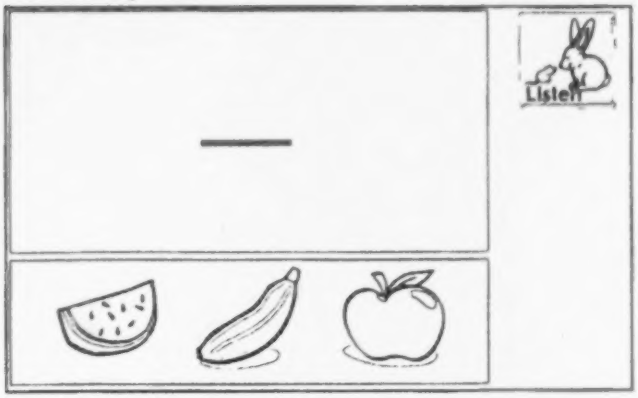
Look at the words: push, child, both.
Click on the word that has the /ch/ sound.

Graphophonemic Knowledge Skills



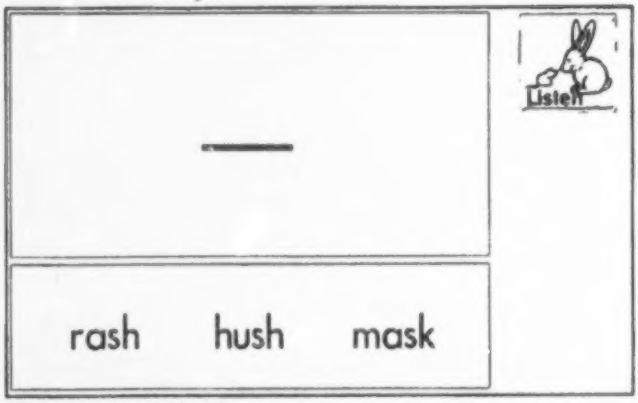
Which of these is the letter v? Click on the letter v.

Vocabulary Skills



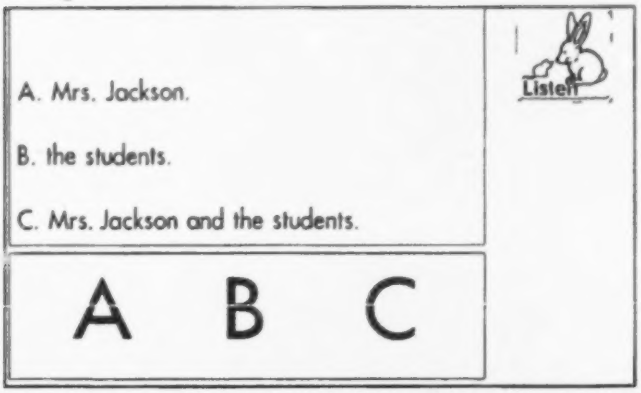
Which picture shows an apple?
Click on the picture of the apple.

Structural Analysis Skills



Listen to this word: ash.
Click on the word you can make from ash.

Comprehension Skills



Listen to the story. Mrs. Jackson liked to read to the class. She read to them almost every day. Her favorite book was about a bear who couldn't fall asleep at night. The students liked this one, too. Now, click on the letter of the answer that tells who liked the story about the bear.

Prototype Research Study
Tryout research of the prototype was carried out in April 2000. Over 1,500 children in pre-kindergarten, kindergarten, and grades one and two participated in the tryout. Extensive analyses were conducted to evaluate the software, its user interface, and the psychometric characteristics and teacher opinions of the test items. The results indicated that the prototype tryout study was a success in terms of demonstrating the viability of the software prototype and of the tryout items in classrooms ranging from pre-kindergarten through grade two. The user interface proved to be usable at all levels, the tasks were well within the ability of children to complete in a minimum of time, the tryout test items demonstrated promising psychometric properties, and teachers generally reacted well to the content and format of the prototype. The few weak points (most were related to correctable audio problems) that were found in the analyses of the tryout study data were addressed in the development of the calibration version of the interface.

Item Calibration: 32,257 Students in 308 Schools Nationwide
In order to use the test items for future adaptive testing, every item had to first be placed on a continuous scale of difficulty; the same scale would be used later to score the adaptive tests. The procedures of item response theory (IRT) were chosen as the basis for scaling STAR Early Literacy item difficulty, a process called "calibration."

IRT calibration is based on statistical analysis of response data—it requires hundreds of responses to every test item. To obtain these data, Renaissance Learning conducted a major item calibration study in late October 2000. For the calibration study, 246 test forms were designed, and 2,900 STAR Early Literacy items[3] were distributed among these forms. Every form contained 40 STAR Early Literacy test items. The forms were graded as to developmental level: Level A forms were designed for pre-kindergartners and kindergartners; Level B was designed for students in first grade; and Level C was designed for use in second grade and above.

Because all STAR Early Literacy test items include audio, these test forms were all computer administered. In November and December 2000, the computer-administered calibration forms were given to a nationwide sample of 32,257 students in pre-kindergarten through grade three, in 308 schools.

Many of the students participating in the calibration study were asked to take two STAR Early Literacy tests, so that the correlation of their scores on two occasions could be used to evaluate the stability of STAR Early Literacy tests over a short time interval.

In addition, a subsample of grade one-through-three students also took the computer-adaptive STAR Reading assessment[4], to provide a basis for evaluating the degree of relationship between STAR Early Literacy and reading ability.

Statistical Analysis: Fitting the Rasch IRT Model to the Calibration Data
With the response data from the calibration study in hand, the first order of business was to calibrate the items and score the students' tests. This was done using the "Rasch Model," an IRT model that expresses the probability of a correct answer as a function of the difference between the difficulty of the item and the ability of the student on a common scale. Rasch Model analysis was used to determine the value of a "difficulty parameter" for every item, and to assign a score to every student. In the course of the analysis, a number of statistical measures of item quality and model fit were calculated for each item.

Selecting Items from the Calibration Item Bank
Once the calibration analysis was complete, a psychometric review took place. Reviewers evaluated each item's difficulty, discriminating power, model fit indices, statistical properties, and content to identify any items that appeared unsuitable for inclusion in the adaptive testing item bank. The review work was aided by the use of interactive psychometric review software developed specifically for STAR Early Literacy.

Of the 2,900 items used in the calibration study, more than 2,500 were accepted for use in the adaptive version of STAR Early Literacy.

Score Scale Definition and Development
Following the completion of item calibration, a score scale was developed for use in reporting STAR Early Literacy results. Although the Rasch Ability Scale could be used for this purpose, a more "user-friendly" scale was preferred.[5] A system of integer numbers ranging from 300 to 900 was chosen as the score reporting scale for STAR Early Literacy.

---

[3] Prior to calibration, 61 items were dropped from the original 2,961 test items.
[4] STAR Reading is a computer-adaptive standardized reading assessment produced by Renaissance Learning, Inc. It contains vocabulary-in-context, authentic text passages, and literal and inferential questions. The assessment contains 1,432 items graded into 54 difficulty levels.
[5] Scores on the Rasch Ability Scale are expressed on the "real number" line, use decimal fractions, and can be either negative or positive. While useful for scientific and technical analysis, the Rasch Ability Scale does not lend itself to comfortable interpretation by teachers and lay persons.

BEST COPY AVAILABLE

Test-Retest Reliability
As mentioned earlier, the calibration study included a test-retest reliability component, in which selected students took calibration tests twice. The two tests were administered on different days, and each student took a different version on retest, to minimize repetition of the same items. The correlation of students' scores on their first and second tests provides one measure of the reliability of STAR Early Literacy tests.[6]

Over 14,000 students took part in the retest reliability study. Figure 1 (on the following page) shows a scatterplot of students' scores on initial test and retest. As the figure indicates, the correlation was very high: .84 overall.

Relationship of STAR Early Literacy Performance to Age and School Grade
The fundamental literacy skills that STAR Early Literacy was designed to measure improve as children mature and as they benefit from instruction. Consequently, if STAR Early Literacy is indeed measuring literacy skills along a developmental continuum, STAR Early Literacy test scores should increase with age and with years of schooling. Table 1 (at the bottom of this page) lists summary statistics for STAR Early Literacy scaled scores by grade.

As these data indicate, scores from the STAR Early Literacy calibration study show the expected pattern of relationship to grade level (and, by implication, to age).

Relationship of STAR Early Literacy Performance to Reading
Besides showing the appropriate relationships with age and grade level, if STAR Early Literacy is indeed measuring literacy skills, its scores should correlate highly with reading measures. To evaluate this, over 3,000 students in grades one through three took STAR Reading tests during the calibration study, in addition to the STAR Early Literacy tests. Figure 2 (on the following page) shows a plot of STAR Early Literacy test scores against their STAR Reading scores. As the shape of the scatterplot suggests, the degree of correlation was substantial: .79 overall.

The STAR Early Literacy Pilot Research Study
The technical results of the STAR Early Literacy calibration study were excellent, with the tests showing good measurement properties, a high degree of reliability, and high correlation with an independent measure of reading ability. However, the calibration study was conducted using conventional tests, while upon release, STAR Early Literacy will be an adaptive test.

The inherent differences between conventional and adaptive test administration raise the possibility that the technical properties of the adaptive version may be somewhat different from those found in the calibration study. Indeed, the adaptive version of STAR Early Literacy is likely to be superior, by virtue of its ability to tailor the choice of test items to the ability level of each student. With that in mind, additional psychometric research data are being collected in the spring of 2001 with a pilot, adaptive version of STAR Early Literacy. Data from this pilot study will be used to assess a number of technical characteristics of the adaptive version, including the following:

- Reliability of the adaptive STAR Early Literacy tests.
- Scale score distributions by age and grade.
- Validity of STAR Early Literacy.
  - External validity: STAR Early Literacy relationships to other tests.
  - Construct validity: Verifying that STAR Early Literacy measures what it purports to measure.
- Appropriateness of the adaptive version of STAR Early Literacy.
  - Mouse and practice item performance.
  - Comparison of actual and target difficulty levels.
  - Test administration time.
- User reactions: Teacher surveys.

Table 1. Summary Statistics for the Calibration Study: STAR Early Literacy Scaled Scores

	Mean	Standard Deviation	Sample Size
Pre-Kindergarten	517	87	2,584
Kindergarten	585	85	5,938
Grade 1	701	83	10,768
Grade 2	763	82	6,852
Grade 3	811	63	6,115

[6] The retest reliability coefficients obtained in the non-adaptive calibration study may be somewhat different from the reliability of the adaptive version of STAR Early Literacy. For that reason, a separate reliability study is being conducted using the adaptive version.

Figure 1: Test-retest Scatterplot of STAR Early Literacy
Initial Test and Retest Scores for 14,230 Students
(Correlation = .84)

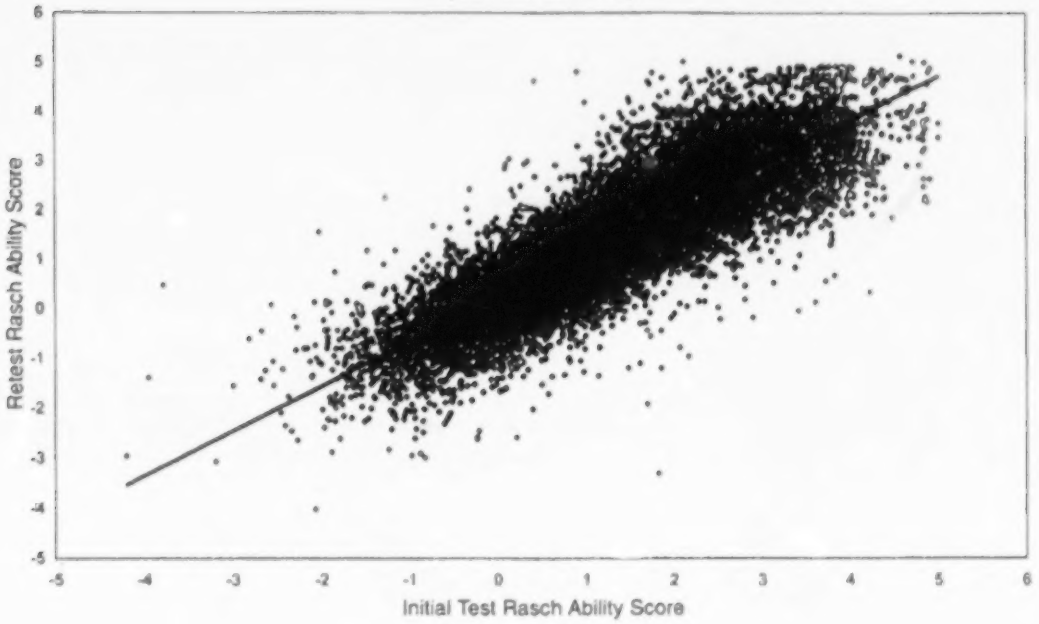
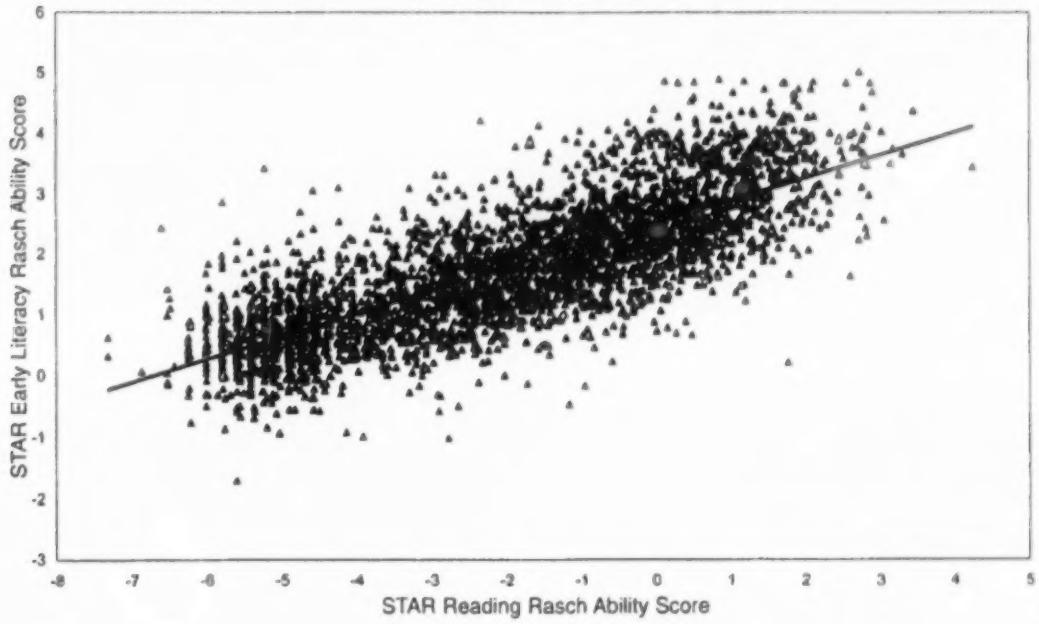


Figure 2: Scatterplot of Rasch Ability Scores from
STAR Early Literacy Calibration and STAR Reading
(Correlation = .79)



Summary

The STAR Early Literacy diagnostic assessment meets the need for an accurate, inexpensive tool to measure the pre-reading skills that are crucial to children's success in reading. It provides educators with relevant, timely information on the development of 41 skills in 7 domains of early literacy skills, enabling more effective and targeted instruction. Its item content was developed in conjunction with leading literacy experts and carefully calibrated using accepted psychometric methods. The user interface was proven through extensive research to be effective with pre-kindergarten through grade three children. In short, STAR Early Literacy promises to be a powerful tool in the hands of educators seeking to improve early literacy instruction.

School Renaissance, STAR Early Literacy, and Renaissance Learning are trademarks of Advantage Learning Systems, Inc. STAR Reading is a registered trademark of Advantage Learning Systems, Inc.

For more information, or for additional copies of this report, contact:

School Renaissance Institute

School Renaissance Institute • Educational Research Department
P.O. Box 45016 • Madison, WI 53744-5016
(800) 200-4848 • www.SchoolRenaissance.com

Creators of Accelerated Reader, Accelerated Math, and the Renaissance Schoolwide Improvement Process

Copyright © 2001 School Renaissance Institute, Inc. All rights reserved.

LD301.0301.RN.10M